

AWS PBS Resources and Examples

This article lists the **qsub** command options that are specific to the HECC AWS Cloud, and provides examples showing how to use them.

Accepted Resource Attributes

You can use these attributes for the **-l select** option:

ncpus=n : minimum n cores per node

mpiprocs=n: n MPI processes per node

mem=nGB : minimum n GB of RAM per node; can be simply nG, ng, or nM, nm.

arch=xx : specifies an architecture type

If not specified, the cheapest instance type that satisfies the user's request will be chosen automatically.

most common ones are: c4/c5/m4/m5
c4 allows for up to 18 cores and 60G per node
c5 allows for up to 48 cores and 192G per node
m4 allows for up to 32 cores and 256G per node
m5 allows for up to 48 cores and 384G per node

gpu=xxx : To specify the GPU node type.
xxx should be either k80 or v100.
The K80 nodes have either 1, 8, or 16 GPUs per node
The V100 nodes have either 1, 4 or 8 GPUs per node

ngpus=n : minimum n gpus per node

net=x : Target network bandwidth in/out of node.
Options supported are 1, 10, 25.
Not specified will allocate a node type based on other specified resource requests only.
If specified will allocate (if possible) a node type that satisfies other requests and:
net=1 -> can sustain around 1+ Gbps,
net=10 -> can sustain 10 Gbps
net=25 -> can sustain 25 Gbps

The following PDFs list the CPU and GPU node types that are supported in AWS:

- [CPU node types](#) (PDF)
- [GPU node types](#) (PDF)

In all regions of AWS, on-demand nodes are used by default. Spot nodes are generally significantly less expensive than on-demand nodes. The drawback is that spot nodes are subject to being taken away based on AWS load and are not always available.

To request spot nodes for your cloud job, use the **ccloud_model** directive in your PBS script:

```
#CLOUD -ccloud_model=[spot, onlspot]:block
```

- Specifying the type "**spot**" will indicate that the job should start using spot instances; if there are not enough spot resources for the job, it will transfer over to an on-demand job type.
- Specifying the type "**onlspot**" will indicate that the job should start using spot instances; if there are not enough spot resources, the job will fail.
- The optional value "**:block**" indicates you want all instances to fail if any single instance is taken away.

For example, the following line will try to start the job on spot instances; if that isn't possible, it will try to start them as on-demand instances. If any of the spot instances is taken away by AWS, the entire job will be terminated:

```
#CLOUD -cloud_model=spot:block
```

On-demand jobs can also fail if resources are not available in AWS. Jobs are automatically retried 10 times before failing due to insufficient resources.

Note: In the PBS script, a **CLOUD** directive should be placed after all PBS directives have been specified.

Job Dependencies

We do not recommend submitting a job (*Job B*) that depends on another job (*Job A*), such as `-W depend=afterany:job_id.server_name.nas.nasa.gov`, because resources for *Job B* may be granted and charging for *Job B* may begin before *Job A* terminates or completes, and before *Job B* actually executes.

Rerunning Jobs

Jobs submitted to run on the AWS cloud cannot be rerun by PBS, no matter what causes the job to fail. Specifying `#PBS -r` has no effect.

Resource Request Examples

- To request 2 nodes with a minimum of 2 cores and 16g of memory:

```
#PBS -l select=2:mpiprocs=2:ncpus=2:mem=16g
```

- To request 3 c4 nodes each with a minimum of 18 cores and 48g of memory:

```
#PBS -l select=3:ncpus=18:mem=48G:arch=c4
```

- To request 2 nodes with at least 1 K80 cards per node:

```
#PBS -l select=2:mpiprocs=2:ncpus=2:mem=16g:gpu=k80:ngpus=1
```

- To request 3 nodes each with a minimum of 14 cores and 4g of memory, plus 2 nodes each with a minimum of 13 cores and 0.5G of memory, and 1 node with at least 3 K80 cards:

```
#PBS -l select=3:mem=4g:ncpus=14+2:mpiprocs=13:mem=512M+1:gpu=k80:mem=16G:ngpus=3
```

- To request 8 nodes, each with 8 V100 GPUs:

```
#PBS -l select=8:mpiprocs=2:ncpus=2:mem=16g:gpu=v100:ngpus=8
```

Article ID: 587

Last updated: 19 Mar, 2021

Revision: 26

Cloud Computing -> AWS Cloud -> AWS PBS Resources and Examples

<https://www.nas.nasa.gov/hecc/support/kb/entry/587/>